

Aus der Medizinischen Universitätsklinik in Marburg
Direktor: Prof. Dr. G. A. Martini

Zuverlässigkeit und Richtigkeit im klinischen Laboratorium

Studie über einen regionalen Ringversuch von Krankenhauslaboratorien

INAUGURAL-DISSERTATION

zur Erlangung des Doktorgrades der gesamten Medizin
dem Fachbereich Humanmedizin

vorgelegt von

Katrin Raschig geb. Hommelhoff
aus Hamburg

Marburg 1975

Aus der Medizinischen Universitätsklinik in Marburg
Direktor: Prof. Dr. G. A. Martini

Zuverlässigkeit und Richtigkeit im klinischen Laboratorium

Studie über einen regionalen Ringversuch von Krankenhauslaboratorien



INAUGURAL-DISSERTATION

**zur Erlangung des Doktorgrades der gesamten Medizin
dem Fachbereich Humanmedizin**

vorgelegt von

**Katrin Raschig geb. Hommelhoff
aus Hamburg**

Marburg 1975



Angenommen vom Fachbereich Humanmedizin
der Universität Marburg am 1.7.1975

gedruckt mit genehmigung des Fachbereichs
Dekan: Prof.Dr.H.Oepen
Referent: Prof.Dr.A.Englhardt
Correferent: Prof.Dr.G.Sturm

Meinem verehrten Lehrer, Frau Prof. Dr. A. Englhardt, danke ich für die Vergabe dieses Themas, die freundliche Anleitung und die stete Förderung, die sie mir während der Arbeit zuteil werden ließ.

Herrn Prof. Dr. Dudeck danke ich für zahlreiche zusätzliche Anregungen zur Auswertung der Ergebnisse des Ringversuches

Der Firma Boehringer Biochemica danke ich für die Bereitstellung und Versendung der Enzym-Kontrollseren an alle Teilnehmer des Ringversuches.

INHALTSVERZEICHNIS

	Seite
1. Einleitung	1
2. Planung und Durchführung	2
2.1. Planung	2
2.2. Durchführung	3
3. Auswertung	4
3.1. Ermittlung der Sollwerte des Kontrollserums	4
3.2. Kriterien zur Beurteilung der Richtigkeit und Präzision	5
3.2.1. Präzision	5
3.2.2. Richtigkeit	5
3.2.3. Varianzanalyse	6
3.2.3.1. Planung	6
3.2.3.2. Statistische Auswertung	6
3.2.3.3. Berechnung der einzelnen Größen	7
4. Ergebnisse	7
4.1. Teilnehmer	7
4.2. Sollwerte des Kontrollserums	8
4.3. Präzision der Messungen	8
4.3.1. Berechnung der relativen Standardab- weichung	8
4.3.2. Varianzanalyse	9
4.3.3. Wechselwirkung Tage - Konzentration	9
4.4. Richtigkeit	10
4.4.1. Auswertung zur Richtigkeit	10
4.4.2. Ergebnisse zu Richtigkeit	11
5. Diskussion	12
6. Tabellen	14
7. Abbildungen	23
8. Zusammenfassung	28
9. Literaturverzeichnis	30

1. EINLEITUNG

Ringversuche auf Anregung und Veranlassung der Fachgesellschaften werden seit längerer Zeit ausgerichtet. Diese Ringversuche ermöglichen eine Beurteilung der Qualität und Richtigkeit der Werte der teilnehmenden Laboratorien (1), (2).

Im Rahmen des Eichgesetzes (Bundesgesetzblatt vom 15. Juli 1969, Teil I, S. 759-770 (3), (4)) werden überregionale Ringversuche durchgeführt, an denen alle Laboratorien teilnehmen sollen, die diagnostische Aufgaben erfüllen.

Der hier beschriebene Ringversuch wurde als Studie auf freiwilliger Basis in einem regional begrenzten Gebiet durchgeführt; er wurde weiterhin begrenzt auf das Teilgebiet der klinischen Chemie Enzyme. Von jedem Laboratorium wurden jedoch mehr Messungen gefordert als in den derzeitigen überregionalen Ringversuchen. Diese Ausweitung und die regionale Beschränkung sollte ermöglichen, verschiedene statistische Auswertungen anzuwenden und zu vergleichen. Ziel der Studie war, zu prüfen mit welchem Auswertungsmodus die Arbeitsweise eines Laboratoriums hinsichtlich Präzision und Richtigkeit seiner Ergebnisse am besten erfaßt werden kann. Darüberhinaus sollte untersucht werden, ob mit diesen Auswertungsmethoden dem Teilnehmer weitergehende Informationen über spezielle Fehler seiner Meßtechnik vermittelt werden können.

Auf unsere Anregung hin wurde in Hessen und angrenzenden Gebieten Nordrhein-Westfalens ein solcher regionaler Ringversuch zur Messung von Enzymaktivitäten gestartet. Eingeladen zur Teilnahme waren alle Laboratorien der Krankenhäuser und Heilstätten dieses Gebietes. Die Auswertung wurde inzwischen abgeschlossen. Sie brachte eine Reihe von interessanten Ergebnissen, über die nachstehend berichtet wird.

Die Studie wurde durchgeführt in enger Zusammenarbeit mit dem Institut für klinische Chemie der Universität Gießen, Direktor Herr Prof. L. Roka. Herrn Prof. Roka sind wir für Anregungen, Mitwirkung und Diskussionen zu großem Dank verpflichtet.

Herrn Prof. G. A. Martini, Direktor der Medizinischen Klinik der Universität Marburg, danken wir für die Förderung dieser Studie.

2. PLANUNG UND DURCHFÜHRUNG

2.1. Planung

Das erste Rundschreiben richtete sich an den Krankenhaussträger in allen Krankenhäusern der erfaßten Gebiete. Es enthielt eine Ankündigung der geplanten Aktion und eine Anfrage nach dem verantwortlichen Leiter des Laboratoriums. Im zweiten Rundschreiben wurden die Leiter der Krankenhauslaboratorien zur Teilnahme am Ringversuch eingeladen.

In der dritten Aussendung erhielten alle, die sich zur Teilnahme ausdrücklich bereit erklärt hatten, Kontrollproben mit genauen Anweisungen zur Handhabung zugesandt.

Zwischenzeitlich wurden folgende Laboratorien als Referenzlaboratorien zu diesem Ringversuch eingeladen, wobei als Auswahlkriterien dienten:

- a) Leitung durch einen anerkannten Klinischen Biochemiker oder Chemiker oder einen Facharzt für Laboratoriumsmedizin.
- b) Regelmäßige Kontrolle der Präzision und Richtigkeit der Enzymmessungen im eigenen Laboratorium durch Qualitätskontrolle.

Laboratorium der 1. Medizinischen Universitätsklinik Düsseldorf (Prof. W. Rick); Institut für Klinische Chemie und Klinische Biochemie am Klinikum Steglitz der Freien Universität Berlin (Prof. K. Borner); Institut für Klinische Chemie an der Universität Gießen (Prof. L. Roka); Laboratorium der Medizinischen Universitätsklinik Marburg/Lahn (Prof. A. Englhardt); Zentrallaboratorium am Krankenhaus Oststadt, Hannover (Prof. A. Delbrück); Zentrallaboratorium des Allgemeinen Krankenhauses Hamburg-Heitberg (Dr. J. Führ); Sektion Klinische Chemie des Zentrums für Innere Medizin und Kinderheilkunde der Universität Elm (Priv. Doz. Dr. Rommel); Hauptlaboratorium der 1. Medizinischen Klinik der Medizinischen Akademie Lübeck (Priv. Doz. Dr. K. Lorentz); Zentrallaboratorium an der 2. Medizinischen Universitätsklinik Kiel (Priv. Doz. Dr. Leybold); Chemisches Institut der Städtischen Krankenanstalten Nürnberg (Prof. G. Hillmann); Medizinisch-Diagnostisches Institut Köln (Dr. D. Laue); Biochemische Abteilung der Firma Merck AG, Darmstadt (Dr. H. Lang); Biochemische Abteilung Biochemica Boehringer, Mannheim.

2.2. Durchführung

Die Teilnehmer erhielten 2 x 10 Fläschchen eines lyophilisierten Kontrollserums mit zwei verschiedenen Konzentrationen von Enzymen. Die zehn Fläschchen der einen Serie mit den höheren Konzentrationen hatten die Bezeichnung "Serie A", die zehn Fläschchen mit den niedrigeren Enzymkonzentrationen hatten die Bezeichnung "Serie B". Als Kontrollserum wurden lyophilisierte in Fläschchen abgefüllte Proben von zwei Chargen eines Kontrollserums der Firma Boehringer (Precinorm) verwendet. Die beiden Chargen hatten verschiedene Konzentrationen von vier Enzymen.

Der Aussendung lagen genaue Anweisungen über das Auflösen der lyophilisierten Proben, Angaben über die Haltbarkeit der Proben und ein Plan zur Durchführung des Versuches bei.

Die Teilnehmer wurden angewiesen, an zehn verschiedenen, nicht unbedingt aufeinanderfolgenden Tagen, jeweils aus einem Fläschchen der Serie A und einem Fläschchen der Serie B zwei Messungen von Enzymaktivitäten durchzuführen. Die Kontrollproben enthielten bekannte Aktivitäten folgender Enzyme:

GOT	Glutamat-Oxalacetat-Transaminase
GPT	Glutamat-Pyruvat-Transaminase
LDH	Lactat-Dehydrogenase
HBDBH	Hydroxybutyrat-Dehydrogenase

Die Laboratorien wurden gebeten:

1. nur diejenigen Enzyme zu messen, die zum Routineprogramm des Laboratoriums gehören,
2. bewährte Routinemethoden zu verwenden und die Ergebnisse in der gewohnten Weise zu berechnen,
3. alle Werte in den der Sendung beiliegenden Ergebnisbogen einzutragen,
4. einen beiliegenden Fragebogen auszufüllen. Dieser enthielt Fragen über die verwendeten Reagenzien und Testvorschriften und die apparative Ausrüstung zur Messung von Enzymen,
5. alle Ergebnisse nur unter der angegebenen Chiffre-Nummer einzusenden.

3. AUSWERTUNG

3.1. Ermittlung der Sollwerte des Kontrollserums

Die Sollwerte des Kontrollserums sind diejenigen Werte, die den wahren Werten so weit wie möglich angenähert sind (2).

Sollwerte wurden auf dreierlei Weise ermittelt:

1. Angaben der Herstellerfirma (Messungen in drei verschiedenen Laboratorien)
2. Berechnung eines mittleren Sollwertes der Messungen von 12 Referenzlaboratorien.⁺

Diese Laboratorien führten an zehn verschiedenen Tagen je zwei Messungen der Serie A und der Serie B von Enzymen durch. Aus den insgesamt 20 Meßwerten der Serie A wurde der Mittelwert $\bar{S}_A$ berechnet, aus den 20 Meßwerten der Serie B der Mittelwert $\bar{S}_B$. Aus allen Werten S_A der Referenzlaboratorien ergab sich der mittlere Sollwert $\bar{S}_A$. Aus allen Werten S_B der Referenzlaboratorien ergab sich der mittlere Sollwert $\bar{S}_B$.

3. Berechnung des Mittelwertes der Messungen aller Teilnehmer. Hierfür wurden zunächst für jedes Laboratorium die Mittelwerte der je 20 Werte der Serie A und B berechnet. Aus diesen ergaben sich die mittleren Sollwerte für A und B (siehe Tabelle 4).

Die Werte $\bar{S}_A$ und $\bar{S}_B$ der Referenzlaboratorien wurden im folgenden bei der Auswertung als Sollwerte zugrunde gelegt.

⁺ Da bei der Ermittlung der Sollwerte der vier Enzyme von den Referenzlaboratorien z.T. zwei voneinander unabhängige Verfahren angewendet wurden und andere Ergebnisse wiederum nicht vollständig waren oder Enzyme nicht gemessen wurden, ergab sich eine unterschiedliche Beteiligung der Referenzlaboratorien an der Ermittlung der einzelnen Sollwerte :

Für GOT nahmen 12 und für GPT 11 Referenzlaboratorien teil, an der Ermittlung der Werte für LDH nur 9, für HBDH nur 7.

3.2. Kriterien zur Beurteilung der Richtigkeit und Präzision(5),(6)

3.2.1. Präzision

Zur Beurteilung der Präzision wurden zwei Verfahren angewendet. Im ersten Verfahren wurden die Berechnungen mit einem Tischcomputer (Olivetti, Programm 101) durchgeführt. Für die vier Messerien der Reihen A_1 , A_2 , B_1 und B_2 wurden Mittelwerte, Standardabweichung und prozentuale Standardabweichung berechnet.

Die zweite Berechnung wurde mit Hilfe einer Varianzanalyse durchgeführt. Diese Berechnung ergab außer der Untersuchung der gesamten Varianz die Teilkomponenten dieser Varianz, nämlich Varianz zwischen den Doppelbestimmungen und Varianz zwischen den Tagen.

Zur Beurteilung der Ergebnisse wurden die Teilnehmer an Hand der Werte für die relative Standardabweichung in vier Gruppen eingeteilt:

Gruppe 1	rel. Standardabweichung	0 - 5%
2		5 - 10%
3		10 - 15%
4		über 15%

Für die Bewertung wurden die Richtlinien der Bundesärztekammer(3) zugrundegelegt. Sie lautet: "Die Präzision wird bei Enzymmessungen als ausreichend betrachtet, wenn die relative Standardabweichung einen Wert von 10% nicht überschreitet". Das heißt, daß nur die Teilnehmer in den Gruppen 1 und 2 mit ausreichend guter Präzision gearbeitet haben.

3.2.2. Richtigkeit

Entsprechend den Richtlinien der Bundesärztekammer(3) werden Werte als richtig bewertet, die innerhalb der dreifachen Standardabweichung der Referenzlaboratorien liegen.

Diese Bewertung erfolgte nach drei verschiedenen Methoden (Tabelle 9):

- I. Durch Prüfung, ob der erste Wert der Reihe A_1 bzw. B_1 innerhalb der dreifachen Standardabweichung der Referenzlaboratorien liegt.
- II. Durch Prüfung, ob die Mittelwerte der Reihen $A_1 + A_2$ und $B_1 + B_2$ innerhalb der dreifachen Standardabweichung der Referenzlaboratorien liegen.
- III. Durch Prüfung, ob die Confidenzbereiche der Mittelwerte der Teilnehmer für die Reihen $A_1 + A_2$ und $B_1 + B_2$ im Bereich der dreifachen Standardabweichung der Referenzlaboratorien liegen.

3.2.3. Varianzanalyse(7)

3.2.3.1. Planung

Im vorliegenden Versuch sollte die Größe von zwei Komponenten des Bestimmungsfehlers von Laboranalysen und die gegenseitige Relation dieser Fehlerkomponenten in den Laboratorien der Teilnehmer des Ringversuches analysiert werden:

1. Fehler zwischen den Doppelbestimmungen (Varianz in der Serie; als Serie wird hier die kleinstmögliche Serie, nämlich zwei Proben, bezeichnet).
2. Fehler zwischen den Bestimmungen an verschiedenen Tagen (Varianz zwischen den Tagen).

3.2.3.2. Statistische Auswertung

Es wird ein varianzanalytisches Modell mit zwei Faktoren angewendet - Zahl der Tage, an denen Messungen durchgeführt werden und Konzentration der zu messenden Proben. Die Messungen an den verschiedenen Tagen werden als Stichproben aus der Gesamtzahl der Tage angesehen, für die die Variabilität bestimmt werden soll, wobei es sich um den "Random-Faktor" handelt. Die Lösungen haben dagegen feste Konzentrationen, d.h. dieser Faktor ist "fixed-Faktor".

Als Modellgleichung ergibt sich:

$$x_{ijk} = \mu + a_i + \beta_j + a\beta_{ij} + \epsilon_{ijk}$$

μ	= Gesamtmittelwert aus allen Proben
a_i	= "Tageseffekt" (Random) ⁺
β_j	= Effekt unterschiedlicher Konzentrationen (fixed)
$a\beta_{ij}$	= Wechselwirkung Tageseffekt-Konzentrationen
ϵ_{ijk}	= Fehler zwischen den Doppelbestimmungen

⁺"Tageseffekt" = Der Anteil der Varianz von Tag zu Tag, die nicht auf der Varianz zwischen den Doppelbestimmungen beruht.

3.2.3.3. Berechnung der einzelnen Größen

A. Varianz zwischen den Doppelbestimmungen

Das mittlere Abweichungsquadrat zwischen den Doppelbestimmungen ist ein Schätzwert für deren Varianz. Diese Varianz gibt die Größe des Fehlers an, der bei der Untersuchung der gleichen Probe am gleichen Tag zu erwarten ist (Präzision innerhalb der Serie).

B. Varianz zwischen den Tagen (Tageseffekt)

Das mittlere Abweichungsquadrat von Tag zu Tag setzt sich zusammen aus der Varianz zwischen den Doppelbestimmungen und der Varianz zwischen den Tagen. Die Varianz $\hat{\sigma}_T^2$ (Varianz zwischen den Tagen) ist ein Schätzwert für die Schwankungen der Tagesmittelwerte, die nicht aus der Varianz zwischen den Doppelbestimmungen erklärt werden können, dies ist ein Schätzwert des echten Tageseffektes". (Ursachen im Labor sind z.B. Temperaturschwankungen unter den Lösungen, Qualität des Wassers, unterschiedliche Reinigung der Laborglaswaren, Verwendung unterschiedlich präziser Pipetten u.a. (8)).

Diesen Wert erhält man nach folgender Formel:

$$\hat{\sigma}_T^2 = \frac{MAQ_T - MAQ_{Do}}{4}$$

Der F-Test MAQ_T / MAQ_{Do} prüft, ob die Varianz gesichert von Null verschieden ist.

4. ERGEBNISSE

4.1. Teilnehmer

Von den in der ersten Aussendung angeschriebenen Krankenhäusern (Tabelle 1) nahm nur etwa ein Drittel schließlich aktiv am Ringversuch teil. Eine Aufstellung der teilnehmenden Krankenhäuser mit Angabe der Bettenzahl und der Art des Krankenhausträgers gibt Tabelle 2, eine Summenübersicht über die für die Arbeit der erfaßten Krankenhauslaboratorien verantwortlichen Leiter enthält Tabelle 3. Diese zeigt, daß die Mehrzahl der Krankenhauslaboratorien des erfaßten Gebietes in Nebentätigkeit von einem Arzt eines anderen Fachgebietes geleitet wird. In vierzig Fällen wurden Angehörige des medizinischen Hilfspersonals als verantwortlich für die Tätigkeit im Laboratorium

angegeben. Nur acht der 137 erfaßten Laboratorien wurden hauptamtlich von Ärzten oder Chemikern mit Erfahrung auf dem Gebiet der Laboratoriumsmedizin geleitet.

4.2. Sollwerte des Kontrollserums

Die drei auf verschiedene Weise ermittelten Sollwerte des Kontrollserums sind in Tabelle 4 aufgeführt.

Bei der Ermittlung des Sollwertes als Mittelwert der Messungen der Referenzlaboratorien und als Mittelwert aller Teilnehmer mußten die angewendeten Testvorschriften berücksichtigt werden. Teilnehmer und Referenzlaboratorien(eins), die mit optimierten Methoden gearbeitet haben, konnten wegen der kleinen Zahl in die Sollwertermittlung und daher auch in die weitere Auswertung leider nicht mit einbezogen werden. Die Werte $\bar{S}_A$ und $\bar{S}_B$ der Referenzlaboratorien wurden im folgenden bei der Auswertung als Sollwerte zugrunde gelegt.

4.3. Präzision der Messungen

4.3.1. Berechnung der relativen Standardabweichung (5)

Zur Beurteilung der Präzision standen vier Meßreihen mit je zehn Werten zur Verfügung, von denen A_1 und A_2 sowie B_1 und B_2 jeweils Doppelbestimmungen derselben Probe waren. Diese wurden jeweils an zehn Tagen durchgeführt, so daß insgesamt 40 Meßwerte von jedem Laboratorium eingesandt wurden. Die relative Standardabweichung konnte daher für vier mal zehn Messungen getrennt, aber auch für alle 40 Messungen gemeinsam berechnet werden.

Tabelle 5 zeigt, welcher Prozentsatz der Teilnehmer mit der zweifachen Standardabweichung im Bereich des $\pm 3s$ -Bereiches der Referenzlaboratorien liegt. Am Hand der Beispiele der Serie A_1 für GOT und LDH (Abb. 1a und 1b) ist die dreifache Standardabweichung der Teilnehmer zum $\pm 3s$ -Bereich der Referenzlaboratorien in Beziehung gesetzt. Wie sich aus Tabelle 5 ergibt, haben nur etwa die Hälfte der Laboratorien Werte, deren zweifache Standardabweichung vollständig im $\pm 3s$ -Bereich der Referenzlaboratorien liegt. Bei Auswertung der dreifachen Standardabweichung der Teilnehmer ist die Zahl noch geringer (siehe Abb. 1a und 1b).

Die Auswertung der Präzision zeigt, daß die Referenzlaboratorien überwiegend in Gruppe 1 und 2 (0-10%) eingestuft werden können (Tabelle 6, Abb. 2a), dagegen ein hoher Anteil der Teilnehmer

nach Gruppe 3 und 4(über 10%) bewertet werden muß(Tabelle 6,Abb.2b); das heißt,die Präzision der letzten Gruppe war nicht ausreichend.

Die Fehler betrafen ausschließlich die Varianz zwischen den Tagen(Tabelle 7);(siehe 4.3.2.).

4.3.2. Varianzanalyse

Die Varianzanalyse ermöglicht die Auftrennung des Gesamtfehlers in die Varianz zwischen den Tagen und die Varianz zwischen den Doppelbestimmungen.Sie ergibt für jedes Laboratorium einen die Präzision charakterisierenden Wert,der sich aus 40 Einzelwerten zusammensetzt.Damit hat die Varianzanalyse einen weiteren Vorteil gegenüber der Berechnung nach Abschnitt 4.3.1..Dadurch ist es möglich,die Präzision der Laboratorien aus allen 40 Werten gemäß den Kriterien der Bundesärztekammer (3) als präzise oder nicht ausreichend präzise einzustufen (Tabelle 6).

Große Varianz zwischen den Tagen wurden bei der Mehrzahl der Teilnehmer mit schlechter Präzision beobachtet.Dagegen waren die Schwankungen zwischen den Doppelbestimmungen ausnehmend gering. Dies spricht dafür,daß bei den Teilnehmern systematische Fehler (5),(6) vorlagen,die weniger in der manuellen Technik als in der technischen Ausrüstung des Laboratoriums zu suchen sind,wie z.B. ungenügende Temperierung,ungenügend präzise Pipetten,ungenügende Reinigung der Glaswaren und anderes mehr(8).

4.3.3. Wechselwirkung Tage - Konzentration

Da an einem Tag zwei verschiedene Konzentrationen gemessen werden, müßten sich systematische Fehler auf alle an einem Tag durchgeführte Messungen in gleicher Weise auswirken.Im vorliegenden Versuchsplan kann außerdem die Wechselwirkung Tage - Konzentration beurteilt werden.

Bei zufälligen Fehlern,wie z.B. Temperaturschwankungen in den Räumen,würden sich bei beiden Konzentrationen gleichgerichtete Verschiebungen nach unten und oben ergeben.In der Varianzanalyse würde sich in diesem Fall keine Wechselwirkung nachweisen lassen.

Eine festgestellte Wechselwirkung würde anzeigen, daß sich die Schwankungen von Tag zu Tag bei den beiden Konzentrationen in unterschiedlichem Ausmaß oder in unterschiedlicher Richtung auswirken. Eine solche signifikante Wechselwirkung würde dann bedeuten, daß nicht systematische Fehler, sondern unkontrollierte Effekte in der Bestimmung die Analysenergebnisse beeinflussen.

Die Grenze der Signifikanz wurde nach Literatur (7) ermittelt.

Diese Wechselwirkungen wurden für alle Laboratorien berechnet und vom Enzym GOT ein nicht signifikantes (Abb. 3a) und ein signifikantes Beispiel (Abb. 3b) graphisch dargestellt. 64% aller Teilnehmer hatten zum Beispiel für GOT signifikante Wechselwirkungen (Tabelle 3), das heißt, die beiden Doppelbestimmungen der Reihen A und B zeigten keine gleichsinnigen Schwankungen an den einzelnen Tagen. Dies bedeutet, daß systematische Fehler in der Meßtechnik keine entscheidende Rolle spielen, sondern andere Effekte, die ohne Kenntnis der Arbeitsweise des Laboratoriums nicht zu erklären sind.

4.4. Richtigkeit

4.4.1. Auswertung zur Richtigkeit (5), (6)

Die Aussage, ob ein Wert richtig ist, ist relativ einfach. Es genügt die Feststellung, daß er im Bereich der dreifachen Standardabweichung von Referenzlaboratorien liegt, deren Werte als richtig zugrundegelegt werden (Tabelle 9).

Eine Sicherung dieser Richtigkeit ist dadurch erreichbar, daß zwei Kollektive von Referenzlaboratorien herangezogen werden, deren mittlere Sollwerte annähernd übereinstimmen müssen. Im vorliegenden Ringversuch wurden die Sollwerte der Referenzlaboratorien und der Herstellerfirma sowie die ermittelten Meßwerte der Teilnehmer miteinander verglichen (Tabelle 4). Die Sollwerte erwiesen sich als ähnlich aber nicht übereinstimmend; die Confidenzbereiche deckten sich nicht vollständig. Dies bedeutet, daß kleine Abweichungen der Teilnehmer vom erlaubten Bereich nicht zu streng bewertet werden dürfen.

In Methode I wurden zwei Einzelwerte zur Beurteilung der Richtigkeit verwendet. Willkürlich wurden die beiden ersten Werte der Reihen A₁ und B₁ herangezogen (Tabelle 9). Diese Methode ergibt nicht immer ein reelles Bild von der Qualität eines Laboratoriums, da diese

Einzelwerte in der Mitte, aber auch am äußeren Rande seiner dreifachen Standardabweichung liegen können. Für den ersten Fall wäre der Einzelwert richtig, im anderen Fall sind aber weitere falsche Einzelwerte zu erwarten, wenn die Präzision schlecht ist.

Ein Beispiel ist das Laboratorium 28, das mit sehr schlechter Präzision arbeitet. Seine Einzelwerte sind einmal richtig, einmal falsch, seine Mittelwerte immer richtig (Abb. 1a). Labor 17 hat sogar mit allen Methoden richtige Werte, obwohl auf Grund der sehr schlechten Präzision falsche Werte zu erwarten sind (Abb. 1a, Tabelle 10 und 11). Methode I ergibt deshalb auch den geringsten Prozentsatz an falschen Resultaten (Tabelle 9).

In Methode II werden die Mittelwerte der Reihe A_1 plus A_2 und B_1 plus B_2 beurteilt. Mit dieser Methode wurden etwas mehr falsche Werte gefunden. Beispielsweise könnte nach Methode I ein Einzelwert noch im erlaubten Bereich liegen, der zugehörige Mittelwert der Serie liegt aber bereits außerhalb dieses Bereiches.

Eine noch kritischere Beurteilung erlaubt die Berechnung der Confidenzbereiche der Mittelwerte der Teilnehmer (Methode III). Sie zeigt, daß wenn der Mittelwert der Reihe noch im erlaubten $\pm 3s$ -Bereich der Standardabweichung liegt, der Vertrauensbereich diesen aber bereits überschreiten kann. Daß die Ergebnisse der Laboratorien je nach Auswertung verschiedener Methoden unterschiedlich ausfallen, wird in Tabelle 10 an 20 Stichproben dargestellt.

4.4.2. Ergebnisse zur Richtigkeit

Das Ergebnis des Ringversuches zeigt, daß bei den Teilnehmern die ausreichende Richtigkeit häufiger ist als die ausreichende Präzision. Nur 24% der Teilnehmer haben sicher falsche Ergebnisse gemessen (Tabelle 9 Methode I). Dieses scheinbar günstige Ergebnis befriedigt jedoch nicht, da, wie die Auswertung der Präzision zeigt, die Meßreihen in großem Umfang falsche Ergebnisse enthalten, die mit den vorliegenden Methoden nicht erfaßt werden können. Daher sollte zur Beurteilung eines Laboratoriums nicht nur die Richtigkeit, sondern auch die Präzision untersucht werden.

5. DISKUSSION

Der vorliegende Ringversuch beschränkt sich bewußt auf ein örtlich begrenztes Gebiet. Es handelt sich überwiegend um Krankenhauslaboratorien, die von Ärzten anderer Fachbereiche oder Angehörigen des Medizinischen Hilfspersonals betreut werden. Besonderer Wert wurde auf eine eingehende Auswertung gelegt, die auf verschiedene Parameter hin durchgeführt wurde. Hierdurch sollten Möglichkeiten und Grenzen der Aussagefähigkeit von Ringversuchen aufgezeigt werden.

Darüberhinaus wurde versucht, durch eine erweiterte Beurteilung Einblick in den Stand der Laboratoriumstechnik der Teilnehmer zu erhalten.

Es ergab sich folgendes :

Im vorliegenden Versuch hatten bei Berechnung der Präzision mit Hilfe der Varianzanalyse für GOT 58% aller Teilnehmer ungenügend präzise Ergebnisse. Der Anteil falscher Werte lag bei 6-24%.

Die Ursachen dieser negativen Quoten konnte nicht eindeutig geklärt werden. Neben dem geringen Anzahl von Teilnehmern beantworteten sämtlichen Fragen über die technische Ausrüstung des Laboratoriums und das Vorgehen bei der Messung. Es fiel auf, daß die Fehler vorwiegend die Varianz von Tag zu Tag und nicht die Varianz zwischen den Doppelbestimmungen betrafen. Dies bedeutet, daß sie weniger auf eine schlechte Technik der ausführenden Kraft als auf eine ungenügende technische Ausrüstung und Organisation des Laboratoriums zurückzuführen sind; z.B. ungenügende Reinigung der Geräte, nicht ausreichende Präzision bei der Herstellung von Reagenzien (8), ungenügend präzise Meßgeräte und Pipetten oder ungenügende Lagerung der Testseren (9). Auffallend war die in hohem Umfang auftretende Wechselwirkung, d.h. die Schwankungen der am selben Tag aus beiden Proben gemessenen Werte waren nicht gleichsinnig. Schwankungen der Raumtemperatur von Tag zu Tag bei Fehlen eines Thermostaten können daher nicht die Ursache der schlechten Präzision sein.

Durch die Erweiterung des Ringversuchs auf zwei Messungen aus je 2 Proben verschiedener Konzentration an 10 Tagen konnten verschiedene Möglichkeiten der Auswertung verglichen werden. Es beeindruckte die Beobachtung, wie schwierig es ist, für ein Laboratorium die eindeutige Aussage 'präzise' oder 'richtig' für die gemessenen Werte zu erstellen.

Je nach der Art der Auswertung wechselt die Anzahl der richtigen und präzisen Ergebnisse aus der Gesamtzahl der Teilnehmer. Dies bedeutet, daß die Präzision und Richtigkeit eines Laboratoriums von Bestimmung zu Bestimmung und von Messerie zu Messerie wechseln kann (Tab. 10), daß aber auch die Art der statistischen Auswertung diese Größe beeinflußt. Es ergibt sich daraus für die Praxis, daß, um zu einer gerechten Beurteilung der Arbeitsweise eines Laboratoriums zu kommen, die Auswertung für alle Ringversuche einheitlich nach demselben Vorgehen erfolgen muß. Es sind nicht nur einheitliche Kriterien (10), sondern auch einheitliche statistische Auswertungsverfahren erforderlich. Größere Meßreihen z.B. Messungen an mehreren Tagen ergeben einen besseren Einblick in die Qualität des Laboratoriums als die Beurteilung von Einzelbestimmungen, da bei diesem Vorgehen diejenigen Laboratorien nicht mit erfaßt werden, deren Werte zufällig trotz schlechter Präzision im Streubereich der Referenzlaboratorien liegen. Auch wird durch die Varianzanalyse eine gute Beurteilung der Präzision ermöglicht. Jedoch hängt es von den personellen und technischen Möglichkeiten der Teilnehmer ab, ob solche Ringversuche realisiert werden können.

Die eindrucksvollen Unterschiede der Präzision der fachlich geleiteten Referenzlaboratorien von der der übrigen Laboratorien ohne fachliche Betreuung zeigt, daß diese von ausschlaggebender Bedeutung ist. Die fachlich geleiteten Laboratorien hatten fast ausschließlich richtige und präzise Ergebnisse. Die fortlaufende Überwachung durch Fachärzte und Klinische Chemiker würde demnach der beste Weg sein, um die Qualität der Ergebnisse nicht nur im Ringversuch wesentlich zu verbessern.

Tabelle 1 Zahl der in den drei Aussendungen erfaßten Laboratorien

1. Aussendungen an die Krankenhausträger	193
2. Aussendungen an die Leiter der Laboratorien mit Einladung zum Ringversuch (durch Antwort der Krankenhausträger ermittelt)	137
nicht erfaßt (keine Antwort des Krankenhaus- trägers)	61
3. Aussendungen an die Teilnehmer mit Kontroll- proben und Anleitung	91
nicht zur Teilnahme bereit	27
Einsendung nicht verwertbarer Ergebnisse	14
Einsendung verwertbarer, vollständiger Ergebnisse	50

Tabelle 2 Die Teilnehmer des Ringversuches (64 Einsender)

A) Einteilung nach der Bettenzahl	
bis 100 Betten	2
100 - 500 Betten	50
500 - 1000 Betten	2
über 1000 Betten	3
B) Einteilung nach der Art des Krankenhausträgers	
Städt. und Kreiskrankenhäuser	34
Konfess. Krankenhäuser	20
DRK - Krankenhäuser	2
Privatkliniken	4
Heilstätten	4

Tabelle 3 Die Leiter der erfaßten und teilnehmenden Laboratorien

	a	b
Chefärzte, Oberärzte und Assistenzärzte folgender Abteilungen :		
Innere Medizin	77	38
Chirurgie	3	2
Hals, Nasen und Ohren	1	-
Pädiatrie	3	2
Augen	1	-
Gynäkologie	2	1
Neurologie	1	1
Apotheker	1	-
Medizinisch-Technische-Assistenten	40	17
Biochemiker, klinische Chemiker und Fachärzte für Laboratoriumsmedizin in hauptamtlicher Tätigkeit	8	3

a im ersten Rundschreiben erfaßte Laboratorien
b am Ringversuch teilnehmende Laboratorien

Tabelle 4 Sollwerte der Chargen A und B des Kontrollserums

Enzym	Charge	Werte in ml / ml		
		a	b	c
GOT	A	52,3	49,2	49,2
	B	36,0	33,6	34,5
GPT	A	53,7	50,5	49,5
	B	31,0	30,4	31,8
LDH	A	207,8	210,6	208,7
	B	155,0	149,2	145,8
HBDH	A	171,7	169,6	155,5
	B	121,0	114,6	110,5

a nach Angaben der Herstellerfirma

b Mittelwerte der Referenzlaboratorien $\bar{S}_A$ und $\bar{S}_B$

c Mittelwerte aller Teilnehmer

(optimierte Testansätze konnten nicht berücksichtigt werden)

Tabelle 5 Prozentsatz der Teilnehmer, deren $\pm 2s$ -Bereich im Vergleich mit dem $\pm 3s$ -Bereich der Sollwerte nach den Kriterien r, a oder f eingestuft wurde.

Enzym	Serie	r	-a	a-	-a-	f
GOT	A ₁	34	20	18	20	8
	A ₂	40	20	16	16	8
	B ₁	38	18	26	14	4
	B ₂	50	16	22	8	4
GPT	A ₁	19	26	26	19	10
	A ₂	17	28	28	17	10
	B ₁	26	15	23	23	13
	B ₂	28	23	21	17	11
LDH	A ₁	58	5	21	11	5
	A ₂	53	11	26	5	5
	B ₁	63	5	11	16	5
	B ₂	53	21	5	16	5
HBDH	A ₁	20	50	10	20	-
	A ₂	30	50	-	20	-
	B ₁	60	10	10	20	-
	B ₂	50	20	10	20	-

- r Bereich der zweifachen Standardabweichung der Meßserie des Teilnehmers liegt innerhalb des Bereiches der dreifachen Standardabweichung der Meßserien der Referenzlaboratorien.
- a Zweifache Standardabweichung der Meßserie überschreitet die dreifache Standardabweichung der Referenzlaboratorien :
nach unten: -a, nach oben: a-, nach oben und unten: -a- .
- f Der Bereich der zweifachen Standardabweichung der Meßserie des Teilnehmers liegt vollständig außerhalb des Bereiches der dreifachen Standardabweichung der Referenzlaboratorien.

Tabelle 6 Prozentsatz der Teilnehmer und Referenzlaboratorien,
deren Präzision an Hand der relativen Standardabweichung
in vier Gruppen eingestuft wurde.

		Teilnehmer ⁺					Referenzlaboratorien ⁺⁺				
		A ₁	A ₂	B ₁	B ₂	A+B	A ₁	A ₂	B ₁	B ₂	A+B
Enzym	Gruppe										
GOT	1	35,0	32,0	26,0	24,0	12,0	55,7	56,7	58,3	75,0	41,6
	2	38,0	44,0	46,0	50,0	30,0	33,3	33,3	33,3	25,0	41,6
	3	10,0	10,0	10,0	12,0	24,0	-	-	8,3	-	16,8
	4	15,0	14,0	18,0	14,0	34,0	-	-	-	-	-
GPT	1	21,3	19,2	17,0	19,2	6,4	63,6	54,5	36,4	45,5	27,2
	2	40,4	38,3	38,3	35,2	17,0	33,4	45,5	54,5	36,4	45,5
	3	25,5	31,9	23,4	27,6	34,0	-	-	9,1	18,2	18,2
	4	12,8	10,6	21,3	17,0	42,6	-	-	-	-	9,1
LDH	1	15,8	10,5	21,1	15,8	10,5	55,6	53,6	33,3	50,6	33,3
	2	47,4	52,6	36,3	42,1	31,6	33,3	33,3	55,6	33,3	33,3
	3	13,8	10,5	15,8	15,8	21,1	-	11,1	-	-	11,1
	4	21,1	26,3	26,3	26,3	36,9	11,1	-	11,1	11,1	-
HBDH	1	20,0	30,0	10,0	30,0	20,0	42,9	28,6	28,6	71,4	14,3
	2	60,0	50,0	60,0	40,0	40,0	57,1	71,4	57,1	28,6	57,1
	3	-	-	10,0	10,0	10,0	-	-	14,3	-	28,6
	4	20,0	20,0	20,0	20,0	30,0	-	-	-	-	-

Gruppe 1 : V% = 0 - 5

Gruppe 2 : V% = 5 - 10

Gruppe 3 : V% = 10 - 15

Gruppe 4 : V% = über 15

GOT GPT LDH HBDH

⁺ Zahl der Teilnehmer :

50 47 19 10

⁺⁺ Zahl der Referenzlab.:

12 11 9 7

Tabelle 7 Übersicht über die Varianz der Messungen berechnet mit einer Varianzanalyse am Beispiel der GOT

rel.Standard- abweichung	0 - 5%		5 - 10%		10 - 15%		über 15%	
Gruppe	1		2		3		4	
Bewertung	Teiln.	Ref.Lab.	Teiln.	Ref.Lab.	Teiln.	Ref.Lab.	Teiln.	Ref.Lab.
A	12,0	41,6	30,0	41,6	24,0	16,6	34,0	..
B	56,0	91,7	30,0	8,3	8,0	-	6,0	-
C	38,0	100,0	12,0	-	-	-	-	-

A Gesamtvarianz MA_{QT}^2

B Varianz zwischen den Tagen (Tageseffekt) $\hat{\sigma}_T^2$

C Varianz zwischen den Doppelbestimmungen MA_{Do}^2

In den Spalten ist der Prozentsatz der Teilnehmer und Referenzlaboratorien angegeben, der auf Grund seiner Präzision in Gruppe 1,2,3 oder 4 einzuordnen ist.

Tabelle 8 Wechselwirkung Tage - Konzentration

Enzym	Anzahl der Teilnehmer in Prozent	
	signifikant	nicht signifikant
GOT	64,0	36,0
GPT	68,1	31,9
LDH	68,4	31,6
HBDH	30,0	70,0

Tabelle 9 Ergebnisse zur Prüfung der Richtigkeit der Werte der Teilnehmer mit drei verschiedenen Methoden am Beispiel des Enzyms GOT
(Anzahl der Teilnehmer angegeben in Prozent)

Methode	Meßreihe	Bewertung		
		richtig	falsch	zweifelhaft
I	A1	86 %	14 %	-
	B1	90 %	10 %	-
II	A1 + A2	76 %	24 %	-
	B1 + B2	88 %	12 %	-
III	A1 + A2	66 %	16 %	18 %
	B1 + B2	74 %	6 %	20 %

Methode I : Berechnung der dreifachen Standardabweichung der Referenzlaboratorien aus dem ersten Wert der Reihen A1 und B1
Kriterium der Richtigkeit: Der erste Wert der Reihe A1 bzw. B1 liegt im Bereich der dreifachen Standardabweichung der Referenzlaboratorien.

Methode II : Berechnung der dreifachen Standardabweichung der Referenzlaboratorien aus den Mittelwerten der Reihen A1+A2 und B1+B2
Kriterium der Richtigkeit: Die Mittelwerte der Reihen A1+A2 und B1+B2 liegen im Bereich der dreifachen Standardabweichung der Referenzlaboratorien.

Methode III: Berechnung der dreifachen Standardabweichung der Referenzlaboratorien wie in Methode II
Kriterium der Richtigkeit: Der Confidenzbereich C des Mittelwertes der Teilnehmer liegt im Bereich der dreifachen Standardabweichung der Referenzlaboratorien.

$$C = \bar{x} \pm t_{0,05} \times \frac{MAQ_{TK}}{n}$$

MAQ_{TK} siehe Tabelle der Varianzanalyse (Tab.12)

Tabelle 10 Ergebnisse zur Prüfung der Richtigkeit am Beispiel der Messung des Enzyms GOT
(20 Stichproben aus den 50 Teilnehmern)

Methode Meßreihe Chiffre- Nummer	I		II		III		IV			
	A1	B1	A1+A2	B1+B2	C _{A1+A2}	C _{B1+B2}	A1	A2	B1	B2
2	f	r	f	r	f	r	f	f	r	r
5	r	r	r	r	z	r	f	f	r	r
9	r	r	r	r	r	r	r	r	r	r
11	r	r	r	r	r	r	r	r	r	r
12	r	r	r	r	r	r	r	r	r	r
13	r	r	f	f	f	z	f	f	f	r
14	r	r	r	r	r	r	r	r	r	r
15	r	r	r	r	r	r	r	r	r	r
17	r	r	r	r	r	r	r	r	r	r
20	r	r	r	r	z	r	f	r	r	r
28	r	f	r	r	r	z	r	r	r	r
29	r	r	r	r	r	r	r	r	r	r
37	r	r	r	r	r	r	r	r	r	r
39	r	r	r	r	r	r	r	r	r	r
41	f	r	r	r	z	z	r	r	r	r
44	r	f	r	f	z	z	r	f	f	r
54	r	r	f	r	z	r	r	r	r	r
61	f	r	r	r	r	r	r	r	r	r
69	f	r	f	f	f	z	f	f	f	r
75	r	r	f	r	f	z	f	f	f	r

Die Richtigkeit wurde mit den Methoden I, II und III (siehe Tabelle 9) beurteilt, zusätzlich wurde jede Meßreihe für sich beurteilt (Methode IV).

r = richtig , f = falsch , z = zweifelhaft

Tabelle 11 Ergebnisse zur Prüfung der Präzision am Beispiel der Messung des Enzyms GOT
(20 Stichproben aus den 50 Teilnehmern)

Chiffre- Nummer	A				B	C	D
2	1	1	1	1	0,99	0,95	0,01
5	2	2	2	2	19,06	0,52	4,63
9	1	1	1	1	2,56	0,90	0,41
11	3	4	2	2	54,27	6,55	11,30
12	1	1	1	2	5,58	0,42	1,29
13	3	3	2	2	25,39	0,07	6,33
14	1	2	2	2	17,18	0,72	4,11
15	1	2	2	1	6,62	1,53	1,27
17	4	3	4	4	152,34	2,55	37,45
20	1	1	1	1	7,56	2,25	1,33
28	4	4	4	4	277,01	5,57	67,86
29	2	2	4	3	52,11	6,05	11,51
37	2	1	1	1	5,63	2,10	0,88
39	1	1	1	1	0,60	0,40	0,05
41	4	4	2	2	141,49	8,25	33,31
44	4	4	4	4	0,03	0,002	0,01
54	2	2	2	2	37,95	1,20	9,19
61	4	4	4	3	229,57	1,70	56,97
69	1	1	1	1	7,37	0,57	1,70
75	1	2	2	3	13,08	2,47	2,65

- A Varianz von Tag zu Tag, getrennt berechnet für die Reihen A_1, A_2, B_1 und B_2 . Angabe der Einstufung in Gruppen wie in Tabelle 16.
- B Varianz von Tag zu Tag, berechnet durch Varianzanalyse für alle 40 Werte der Reihen $A_1 + A_2 + B_1 + B_2$, angegeben als MAQ_m (siehe Tabelle der Varianzanalyse Tab. 12). (Grenzwert 17,20)
- C Varianz der Doppelbestimmungen, berechnet durch Varianzanalyse wie unter B, angegeben als MAQ_{DO} (siehe Tabelle 12)
- D Varianz zwischen den Tagen "Tageseffekt", berechnet wie unter B.

Tabelle 12: Tabelle der Varianzanalyse

Varianz	Summe der Abweichungs- quadrate (SAQ)	f	mittleres Abweichungs- quadrat (MAQ)	Erwartungs- wert	F-Test
von Tag zu Tag	$4 \sum_{j=1}^{10} (\bar{x}_j - \bar{x})^2$	9	$MAQ_T = \frac{4}{9} \sum_{j=1}^{10} (\bar{x}_j - \bar{x})^2$	$\sigma^2 + 4\sigma_T^2$	MAQ_T / MAQ_{Do}
zwischen den Konzentrat.	$20 \sum_{i=1}^2 (\bar{x}_i - \bar{x})^2$	1	$MAQ_K = 20 \sum_{i=1}^2 (\bar{x}_i - \bar{x})^2$	$\sigma^2 + 2\sigma_{TK}^2 + 10\sigma_K^2$	$\frac{MAQ_K}{MAQ_{TK}}$
Wechselwirk. Tage-Konzent.	$2 \sum_{i=1}^2 \sum_{j=1}^{10} (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})^2$	9	$MAQ_{TK} = \frac{2}{9} \sum_{i=1}^2 \sum_{j=1}^{10} (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})^2$	$\sigma^2 + 2\sigma_{TK}^2$	MAQ_{TK} / MAQ_{Do}
zwischen den Doppelbest.	$\sum_{i=1}^2 \sum_{j=1}^{10} \sum_{k=1}^2 (\bar{x}_{ijk} - \bar{x}_{ij})^2$	20	$MAQ_{Do} = \frac{1}{20} \sum_{i=1}^2 \sum_{j=1}^{10} \sum_{k=1}^2 (\bar{x}_{ijk} - \bar{x}_{ij})^2$	σ^2	
Total	$\sum_{i=1}^2 \sum_{j=1}^{10} \sum_{k=1}^2 (\bar{x}_{ijk} - \bar{x})^2$	39			

$\bar{x}$ Gesamtmittelwert

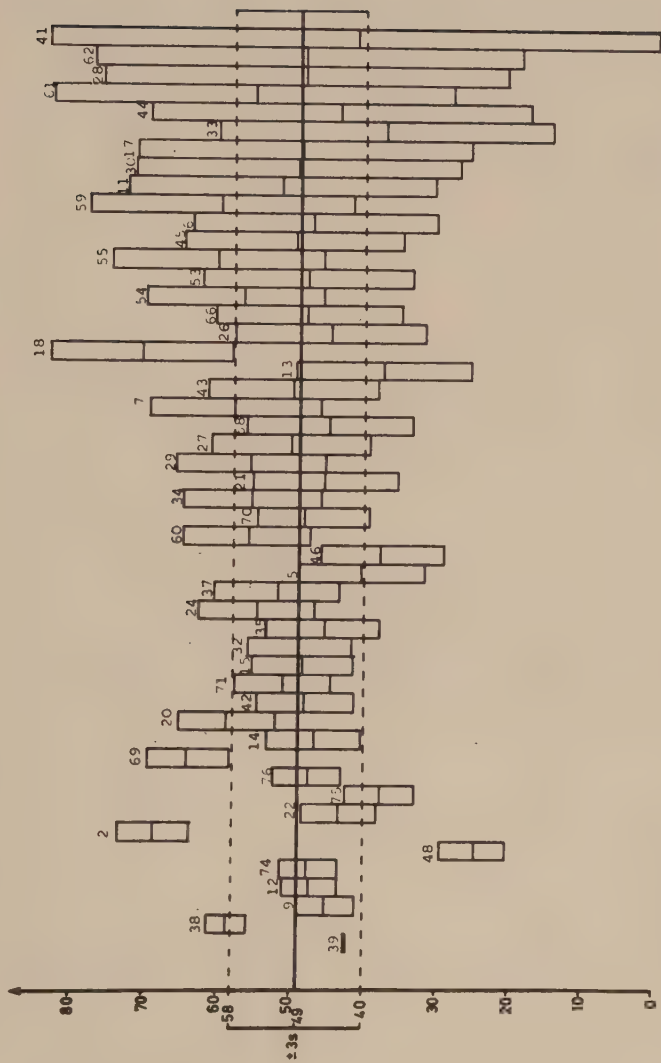
$\bar{x}_j$ Tagesmittelwert

$\bar{x}_i$ Mittelwert bei den verschiedenen Konzentrationen

$\bar{x}_{ij}$ Mittelwert für die Konzentration i am Tag j

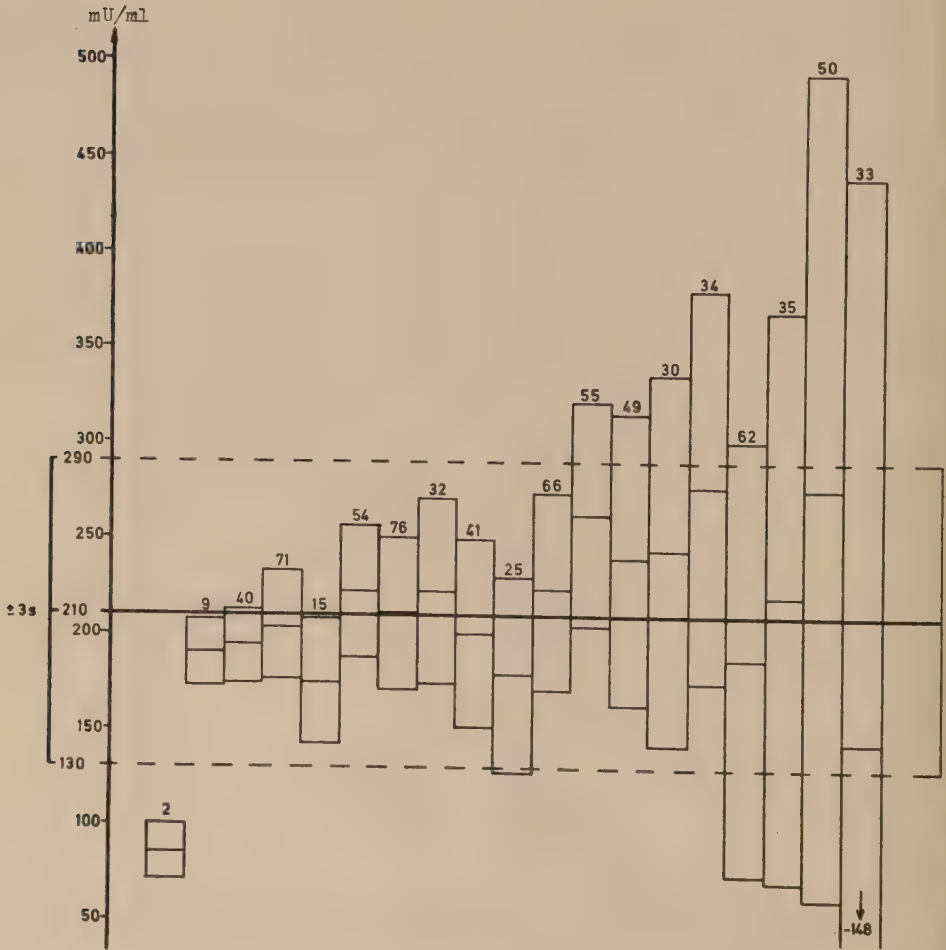
$\bar{x}_{ijk}$ Einzelmeßwert

Abb. 1a Vergleich der Mittelwerte ($\pm 3s$ -Bereiche) der Teilnehmer mit dem Sollwert ($\pm 3s$ -Bereich); Serie A₁ der GOT



Die Ziffern stellen die Chiffre-Nummern der Teilnehmer dar.

Abb. 1b Vergleich der Mittelwerte ($\pm 3s$ -Bereiche) der Teilnehmer mit dem Sollwert ($\pm 3s$ -Bereich); Serie A₁ der LDH



Die Ziffern stellen die Chiffrenummern der Teilnehmer dar.

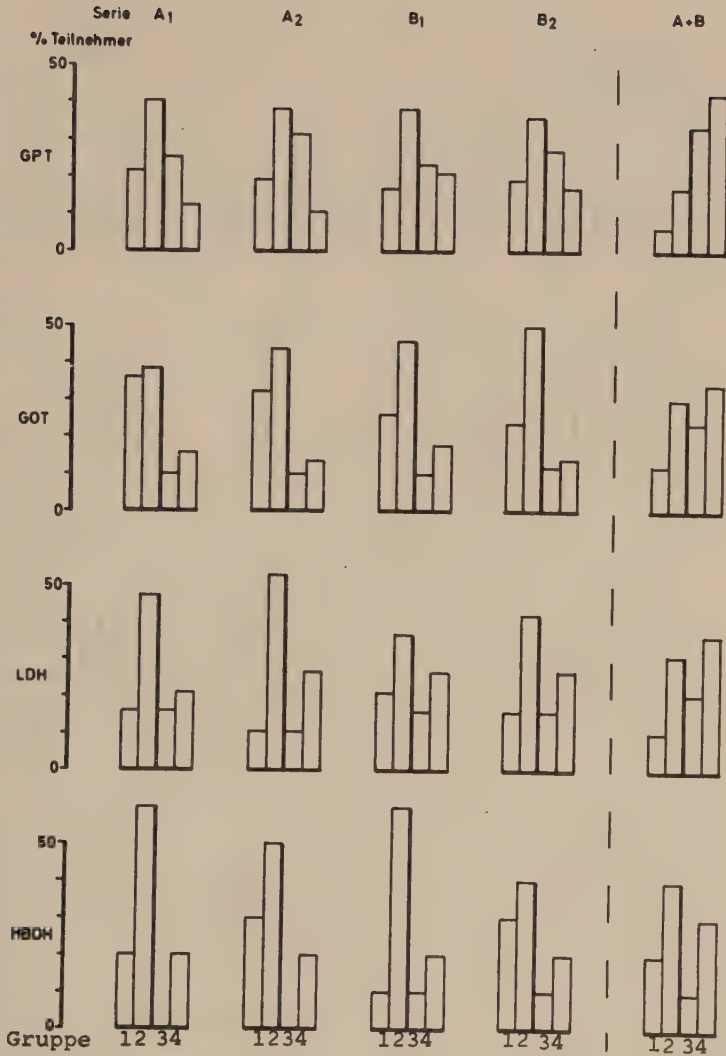


Abb.2b Präzision der Teilnehmer an Hand der relativen Standardabweichung

Gruppe 1: V% = 0 - 5%
 Gruppe 2: V% = 5 - 10%

Gruppe 3: V% = 10 - 15%
 Gruppe 4: V% = über 15%



Abb.2a Präzision der Referenzlaboratorien an Hand der relativen Standardabweichung

Gruppe 1: V% = 0 - 5%
 Gruppe 2: V% = 5 - 10%

Gruppe 3: V% = 10 - 15%
 Gruppe 4: V% = über 15%

Differenz zwischen den einzelnen Tagesmittelwerten für das Enzym GOT bei verschiedenen Konzentrationen: — Konzentration A, --- Konzentration B

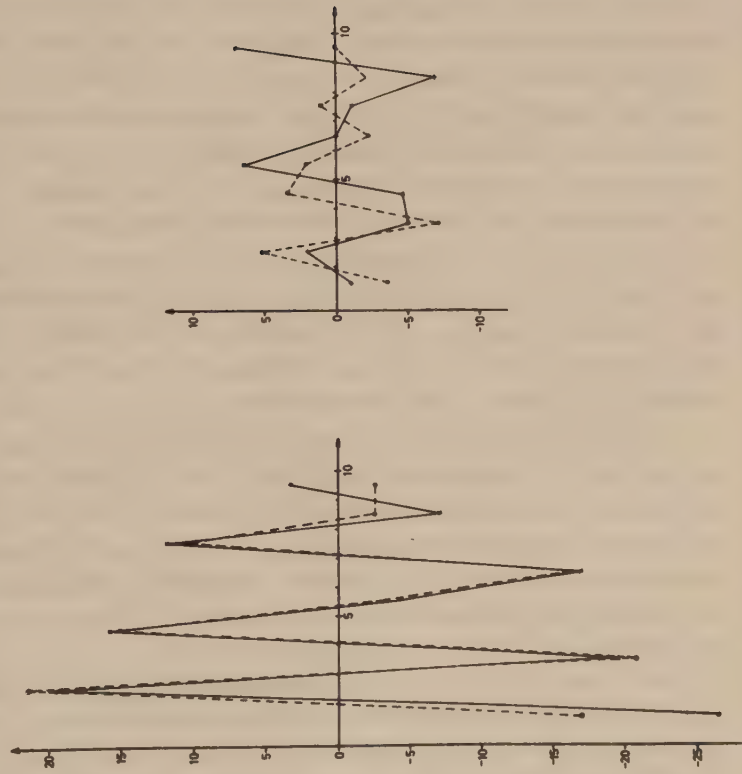


Abb. 3a: nicht signifikant $F=0,71$ Abb. 3b: signifikant $F=208,7$
Chiff.-Nr. 44 Chiff.-Nr. 13

8. ZUSAMMENFASSUNG

Es wird über die Ergebnisse eines regional durchgeführten Ringversuches berichtet, welcher die Beurteilung der Richtigkeit und der Präzision bei Enzymbestimmungen (GOT, GPT, LDH, HBDH) ermöglicht. Mit 64 Laboratorien beteiligten sich schließlich nur etwa ein Drittel aller zunächst erfaßten Krankenhäuser und Heilstätten dieser Region. Die Referenzwerte wurden aus den Werten von 12 anerkannten, fachlich geleiteten Laboratorien ermittelt, die aus dem gesamten Bundesgebiet waren. Zwei Kontrollseren unterschiedlicher Enzymkonzentration wurden laut Versuchsplan an zehn nicht unbedingt aufeinanderfolgenden Tagen jeweils als Doppelbestimmung vermessen.

Neben der konventionellen Methode zur Berechnung der Präzision wurde eine Varianzanalyse durchgeführt, welche die Berechnung der Gesamtvarianz sowie deren Teilkomponenten Varianz von Tag zu Tag und Varianz zwischen den Doppelbestimmungen ermöglichte.

Zur Beurteilung der Richtigkeit wurden einmal zufällig ausgewählte Werte der Meßreihen des Teilnehmers zum anderen die Mittelwerte aller Messungen jeweils einer Enzymkonzentration und schließlich die "Confidenzbereiche" der Mittelwerte der Teilnehmer für die Reihen $A_1 + A_2$ und $B_1 + B_2$ herangezogen.

Die Prüfung auf Richtigkeit bei den Teilnehmern beispielsweise für die höhere Konzentration A des Enzyms GOT ergab bei der Beurteilung eines zufällig ausgewählten Meßwertes bei 86% der Teilnehmer, beim Mittelwert der Meßreihe bei 76% und beim Confidenzbereich des Mittelwertes bei 66% der Teilnehmer richtige Ergebnisse, d.h. die jeweils betrachteten Werte lagen im $\pm 3s$ -Bereich des Sollwertes.

Betrachtet man nicht Einzel- bzw. Mittelwerte für sich alleine unter diesem Kriterium der Richtigkeit, sondern zusätzlich die $\pm 2s$ -Bereiche der Labormittelwerte, so ermöglicht dies eine erweiterte Beurteilung. Unter diesem erweiterten Kriterium lieferten nur knapp die Hälfte aller Teilnehmer Werte, deren $\pm 2s$ -Bereich vollständig im $\pm 3s$ -Bereich der Referenzlaboratorien lag. Die übrigen überschritten diesen Bereich nach einer oder nach beiden Richtungen. 4-13% lagen mit allen Werten außerhalb dieses $3s$ -Bereiches.

Bei der Auswertung der Präzision zeigte sich, daß die Referenzlaboratorien fast ohne Ausnahme relative Standardabweichungen der Werte bis zu 5% bzw. 10% aufwiesen, die Teilnehmer aber in hohem Ausmaß relative Standardabweichungen bis zu 15% oder gar darüber erreichten; das bedeutet, daß die Präzision dieser Teilnehmer nicht den Richtlinien der Bundesärztekammer entsprach.

Die Varianzanalyse der Werte, wiederum am Beispiel der GOT, zeigte, daß die relativen Standardabweichungen zwischen Doppelbestimmungen jedes Tages bei allen Teilnehmern bei zulässigen Werten bis zu 10% lagen, wohingegen bei der Varianz zwischen den Tagen relative Standardabweichungen bis zu 10% nur von 86% der Teilnehmer erreicht wurde. Die Gesamtvarianz erbrachte nur bei 42% der Teilnehmer relative Standardabweichungen bis zu 10%.

Aus diesen Ergebnissen ist zu schließen, daß die schlechte Präzision durch hohe Varianz zwischen den Tagen und nicht zwischen den Doppelbestimmungen bedingt war. Dies spricht mehr für eine ungenügende technische Ausrüstung des Laboratoriums als für eine fehlerhafte manuelle Technik.

Die Beurteilung der Richtigkeit der Ergebnisse zeigte häufiger eine ausreichende Richtigkeit als eine ausreichende Präzision. Dies ist dadurch bedingt, daß mit Hilfe der eingesandten Stichproben nicht alle falschen Ergebnisse, die die Teilnehmer mit schlechter Präzision liefern würden, erfaßbar sind.

Die Ursache der hohen Fehlerquoten ließ sich weder durch die angewandten Auswertungsmethoden, noch durch Fragebogen ermitteln. Die hohe Qualität der Ergebnisse der Referenzlaboratorien zeigt aber, daß präzise und richtige Ergebnisse am besten durch dauernde Fachaufsicht erreicht werden können.

9. Literaturverzeichnis

- (1) von Boroviczény,K.G. ,Merten,R.
Ärztl.Lab. 18, 58-74 (1972)
- (2) Merten,R
Ärztl.Lab. 17, 30-34 (1971)
- (3) Deutsches Ärzteblatt 68, 2228-2231 (1971)
- (4) von Boroviczény,K.G.
Ärztl.Lab. 17, St 1-4 (1971)
- (5) Büttner,H. ,Stamm,D.
Z.Klin.Chem. 4, 303-306 (1966)
- (6) von Boroviczény,K.G. ,Merten,R.
Ärztl.Lab. 17, 384-391 (1971)
- (7) Scheffé, Henry : The Analysis of Variance
New York 1964
- (8) Z.Klin.Chem.Klin.Biochem. 8, 658-660 (1970)
- (9) Szasz,G.
Z.Klin.Chem.Klin.Biochem. 8, 212-217 (1970)
- (10) Stamm,D.
Z.Klin.Chem.Klin.Biochem. 12, 137-145 (1974)

Lebenslauf

Ich, Katrin Raschig, geb. Hommelhoff, wurde am 14.10.1945 in Hamburg als Tochter des praktischen Arztes Dr.med. Peter Hommelhoff und seiner Ehefrau Dr.med. Lotte Hommelhoff, geb. Sieger, geboren.

1952 wurde ich in die Jahnschule in Hamburg eingeschult. Von 1956-1964 besuchte ich das neusprachliche Gymnasium der Heilwig-Schule in Hamburg und wechselte 1964 auf das private Landerziehungsheim Stein/bei Traunstein über. 1966 legte ich als Privatschülerin des Maximilian-Gymnasiums in München die Reifeprüfung ab.

Im Wintersemester 1965/67 studierte ich Chemie in Würzburg. Danach begann ich in München mit dem Studium der Humanmedizin, wo ich im April 1968 das Vorphysikum und im Oktober 1969 das Physikum bestand. Vom Wintersemester 1969/70 an setzte ich mein Studium in Marburg an der Lahn fort und legte dort im Februar 1973 das medizinische Staatsexamen ab.

Meine akademischen Lehrer waren:

Autrum, Bachmann, Beckmann, Berendes, Bommer, Buchholz, Bücher, Dane, Dannenberg, Daume, Dingler, Dölle, Dziallas, Engelhardt, Exner, Faessler, Frick, Friederich, Graul, Hamelmann, Hamm, Hardewig, Hess, Heuser, Hildebrandt, Hort, Huisgen, Hupe, Jacob, Kaffarnik, Koransky, Kramer, Kranz, Linneweh, Mann, Martini, Merxmüller, Pots, Reim, Rodeck, Rohen, Schleyer, Schmid, Schmitz-Moormann, Seidel, Siegert, Staub, Strohmeyer, Voß, Wallraff, Wiberg.

